

Effectiveness of Using the Modified Checklist for Autism in Toddlers in Two-Stage Screening of Autism Spectrum Disorder at the 18-Month Health Check-Up in Japan

Yoko Kamio · Naoko Inada · Tomonori Koyama · Eiko Inokuchi · Kenji Tsuchiya · Miho Kuroda

Published online: 6 June 2013
© Springer Science+Business Media New York 2013

Abstract To determine whether the Modified Checklist for Autism in Toddlers (M-CHAT) in conjunction with the routine 18-month health check-up identifies Japanese toddlers with autism spectrum disorder (ASD). Two-stage screening using the M-CHAT was conducted with 1,851 children attending the check-up. Final ASD diagnosis was confirmed at age ≥ 3 years. Screening identified 20/51 children with ASD: 12/20 true positives were developmentally delayed, whereas 16/22 false negatives were high-functioning. Sensitivity was 0.476, specificity 0.986, positive predictive value 0.455, and likelihood ratio 33.4 for children with ASD. With a few modifications, M-CHAT screening successfully detected toddlers with ASD with and without developmental delay and is a promising

screening tool to complement existing community surveillance.

Keywords Autism spectrum disorder · Community-based surveillance · Early detection · Modified Checklist for Autism in Toddlers (M-CHAT) · Screening

Introduction

Autism spectrum disorders (ASDs) are lifelong developmental disorders and the earliest symptoms start to manifest overtly from the age of 1 year onwards. Since early educational intervention can optimize long-term prognosis (Kamio et al. 2013; Rogers and Vismara 2008), early detection and diagnosis are crucial. The American Academy of Pediatrics (AAP) recommends that in addition to broad developmental screening at 9, 18, and 24 months, all children receive autism-specific screening at 18 and 24 months of age, and it cautions against a “wait-and-see” approach for children with suspected ASD (Johnson and Myers 2007). Although many screening tools are available for children aged 18 months and older (Johnson and Myers 2007), several issues such as the optimal age for screening, general developmental surveillance versus standardized autism-specific screening, and barriers to standardized screening remain to be answered by a series of longitudinal studies (Barton et al. 2008; Charman et al. 2001). Moreover, most screening tools have been evaluated in clinical samples referred for specialized assessment (Allen et al. 2007; Eaves et al. 2006) or in a mixture of clinical and population-based samples (Robins et al. 2001); only a few have been examined in total population studies (Baird et al. 2000; Dietz et al. 2006; Pandey et al. 2008; Robins 2008). Also, parents who do not suspect their child to have ASD

Y. Kamio (✉) · N. Inada · T. Koyama · E. Inokuchi · M. Kuroda
Department of Child and Adolescent Mental Health, National
Institute of Mental Health, National Center of Neurology and
Psychiatry, 4-1-1 Ogawa-Higashi, Kodaira, Tokyo 187-8553,
Japan

e-mail: kamio@ncnp.go.jp

N. Inada
e-mail: nainada-ky@umin.ac.jp

T. Koyama
e-mail: tomok-ky@umin.ac.jp

E. Inokuchi
e-mail: eiko_i@osa.att.ne.jp

M. Kuroda
e-mail: pr6m-krd@asahi-net.or.jp

K. Tsuchiya
Research Center for Child Mental Development, United
Graduate School of Child Development, School of Medicine,
Hamamatsu University, Hamamatsu, Japan
e-mail: tsuchiya@hama-med.ac.jp

may respond to the same screening questions differently from those who do suspect it, and the results of screening should be interpreted cautiously if screening tools are used outside the setting in which their psychometric properties are known to apply (Gray et al. 2008).

Among the autism screening tools available, the Checklist for Autism in Toddlers (CHAT) (Baron-Cohen et al. 1992) was the first. In a total population study ($n = 16,235$) with follow-up from age 18 months up to 7 years (Baird et al. 2000), two-stage CHAT screening of 18-month-old children identified 10 of 94 children with Pervasive Developmental Disorders (PDDs) using the high-risk threshold, showing a sensitivity of 0.106, a specificity of 1.00, and a positive predictive value (PPV) of 0.833. In another study, two-stage screening of 31,724 children aged 14–15 months using the Early Screening of Autistic Traits Questionnaire (ESAT) identified 18 children who were diagnosed with ASD at an average age of 23.3 months, giving a PPV of 0.25 (Dietz et al. 2006). The Modified Checklist for Autism in Toddlers (M-CHAT) was developed as a more sensitive alternative to the CHAT (Robins et al. 2001) and has been extensively validated (Chlebowski et al. 2013; Pandey et al. 2008; Robins 2008; Kleinman et al. 2008), although its psychometric properties confirmed through long-term follow-up were determined for a combined clinical and low-risk sample (Kleinman et al. 2008). Against this background, the present study evaluated the utility of M-CHAT screening for Japanese toddlers in primary health settings. We targeted children aged 18 months for practical reasons: all Japanese children have a regular general health check-up at 18 months of age, as stipulated by the Maternal and Child Health Act, and the attendance rate is over 90 % (Mothers' & Children's Health & Welfare Association 2007).

Methods

Catchment Area

The catchment area was the suburbs of Fukuoka City, one of the biggest cities in Japan. Its total population is 93,093 according to the 2003 administrative register. The 2000 national census shows that 74 % of the working population is employed in manufacturing with the remainder working in the commerce, service, agriculture, forestry, or fishery sectors.

Participants

From April 2004 to March 2007, 2,141 children (95.4 % of the 2,245 total population cohort) attended the routine 18-month health check-up at a local health center. Written informed consent to participate in this study was obtained

from the parents of 2,113 children (consent rate = 98.7 %). Exclusion of 262 children without any follow-up data after age 3 left 1,851 children (87.6 %) for the subsequent analyses (Table 1). The 262 children excluded and the remaining 1,851 children were not significantly different in terms of sex ratio, mean age at M-CHAT screening, or screening results.

Screening Tool

Children were screened using the Japanese version of the Modified Checklist for Autism in Toddlers (M-CHAT-JV). Its high mother-father and test-retest reliability as well as concurrent and discriminant validity for Japanese toddlers have been reported (Inada et al. 2011). The majority of the Japanese general population aged 18 months has been confirmed to manifest all of the preverbal social behaviors screened by the M-CHAT-JV (Inada et al. 2010).

Because the original M-CHAT was intended to target children aged 2 (Robins et al. 2001), we assumed that the threshold might miss some children aged 18 months in a non-selected population. A preliminary analysis of data from the first one hundred 18-month-old children showed that the total 3 criteria used in the original study (Robins et al. 2001) still worked to identify possible cases ($n = 7$), but the critical 2 criteria identified only one in 100 children and missed 6 of 7 possible cases. In light of this, we modified the original threshold by defining 10 items as our key item set (comprising the original 6 items and newly added items 6, 20, 21, and 23) and lowered the threshold for the first-stage screening by replacing the original first-stage threshold of “any 3 from the total 23 or any 2 from the critical set criteria” with “any 3 from the total 23 or any 1 from the critical set criteria”. For the second-stage screening, we adopted the original threshold, namely a total of 3 or any 2 from the critical set criteria.

Procedure: Screening and Follow-Up

1. *Screening using the M-CHAT* (Fig. 2) Our two-stage screening consisted at the first stage of administering the M-CHAT-JV at 18 months of age (any 3 from the total 23 or any 1 from the critical set criteria) and at the second stage of a follow-up telephone interview (FUI) at 19–20 months of age (any 3 from the total 23 or any 2 from the critical set criteria). The FUI followed a translated script with specific examples in which all failed items were reviewed with a parent in accordance with the original procedure (Robins et al. 2001). When reviewing the failed responses with the parents, trained interviewers did not use the term ‘fail’ and attempted not to cause anxiety or distress for the parents. They also offered feedback or advice when necessary. Parents were provided concrete examples of the target behaviors

Table 1 Characteristics of participants

	Participants classified as having ASD (n = 51) n (%), mean (SD), range	Total participants (n = 1851) n (%), mean (SD), range
Sex ratio, M : F	35 : 16	942 : 909
Age at M-CHAT-JV (months)	18.6 (0.6) 18–21	18.7 (0.6) 17–26
M-CHAT-JV total (failed items)	4.1 (3.2) 0–13	1.0 (1.4) 0–13
M-CHAT-JV critical 10 (failed items)	2.3 (2.2) 0–8	0.3 (0.9) 0–8
Number of evaluations	1.9 (0.8) 1–3	-
Age at final evaluation	50.6 (14.2) 33–73	-
IQ/DQ ^a	80.1 (26.7) 20–134	-
Developmental delay ^b	26 (51.0 %)	-
Participants diagnosed with ASD by the research team (n = 34)		
AD : other ASD, (boys)	16 (14) : 18 (11)	
No. of evaluations	2.3 (0.6) 1–3	
Age at final evaluation (months)	49.4 (11.5) 33–73	
CARS total scores	34.0 (4.7) 24.5–44.5	
ADI-R toddler total scores ^c	25.5 (7.5) 11–39	
ADOS (a) + (b) total scores ^d	13.4 (3.8) 9–23	
IQ/DQ	82.1 (28.1) 20–134	
≥85	17	
70–84	4	
50–69	8	
35–49	4	
<35	1	

AD autistic disorder, ADI-R the Autism Diagnostic Interview-Revised, ADOS the Autism Diagnostic Observation Schedule, ASD autism spectrum disorder, CARS the Childhood Autism Rating Scale, M-CHAT-JV the Japanese version of the Modified Checklist for Autism in Toddlers

^a 43 of 51 participants were assessed by standardized intellectual/developmental tests

^b In addition to the 43 participants with IQ/DQ data, 8 participants were clinically judged on the presence of developmental delay

^c 30 participants were evaluated using the ADI-R

^d 19 participants were evaluated using the ADOS

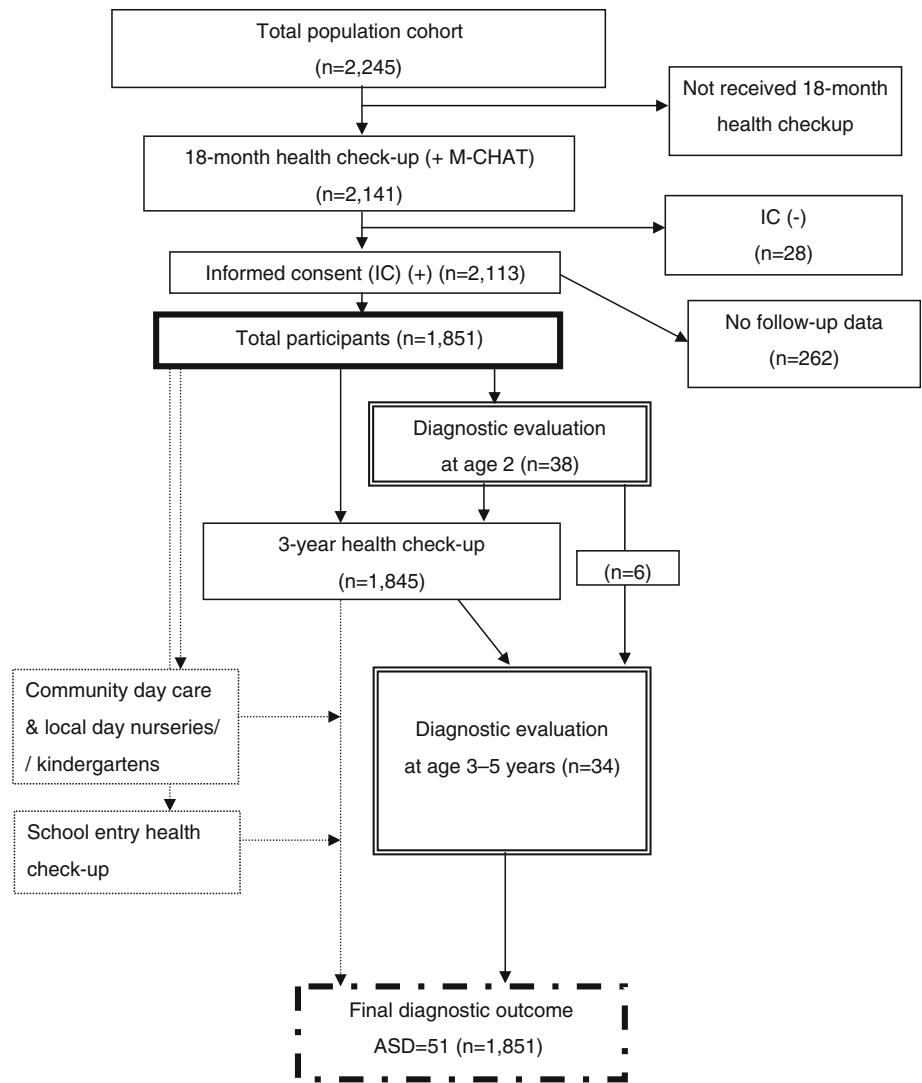
in order to help our judgment of their responses. If the child continued to fail the M-CHAT-JV after the FUI, the family was told that their child was not doing some things that were important for social communication at this age and an evaluation was recommended (Fig. 1).

2. *Diagnostic evaluation at age 2* Screen positives were invited for diagnostic evaluation at age 2. Evaluations were conducted by the research team consisting of child psychiatrists, licensed psychologists, and primary care nurses who were already familiar with the children with special needs. The evaluation instruments included the Japanese versions of the Childhood Autism Rating Scale (CARS) (Kurita et al. 1989; Schopler et al. 1988), the Autism Diagnostic Interview-Revised (ADI-R) (Tsuchiya et al. 2012; Lord et al. 2000), and the Autism Diagnostic Observation Schedule (ADOS) (Lord et al. 1994). Children who were evaluated at age 2 were invited for full evaluation at ages 3, 4, and 5, irrespective of the diagnosis at this age.

3. *Routine 3-year health check-up* Children at age 3 received a routine health check-up including pediatric examination and parental interview by primary care nurses. Parental interviews were conducted based on a checklist containing autism-specific items derived from the ADI-R. The items included in the checklist comprised 10 social domain items, 8 communication domain items, and 2 repetitive or restricted behavior items. Among the 20 items in total, 7 items were picked up from the conventional checklist used for the routine health check-up at age 3 and 13 items were modified from the ADI-R items and newly added.

The social domain items inquire about eye contact, facial expression, nodding as yes, interest in peers, attracting adults' attention, point following, showing as joint attention, play with mother, play with peers, and social reference. The communication domain items ask about imitating what mother does, pretend play by himself/

Fig. 1 Study design



herself, pretend play with others, saying only words, saying his/her name, speaking 2-word sentences, understanding what he/she is said, and using why or what questions. The repetitive or restricted behavior domain items ask about being upset when a routine is broken or when in new environment, and stereotyped movement.

In a pilot study of 39 consecutive children who received the 3-year health check-up, failing more than 3 social or communication items produced a sensitivity of 0.857 and a specificity of 0.400 (Kamio et al. unpublished). Therefore, in the present study, this threshold in combination with behavioral observation by the primary health professional was used to detect false negative children at age 3. Among 1,830 children whose item records had no missing data, 2.24 % (41/1,830) failed more than 3 items, suggesting that the second screening at age 3 may be helpful for detecting false negatives.

The 20-item autism-specific checklist used was created in order to follow up as many false negatives as we could at

age 3. That is, children who were suspected of having ASD at age 3 based on the parental interview using the checklist or on behavioral observation during the medical examination were invited, along with screen-positive children, for full follow-up evaluation including the CARS, ADI-R, or ADOS at ages 3, 4, and 5.

4. *Community day care and local day nurseries/kindergartens* More than 90 % of the participating children went to local day nurseries or kindergartens during preschoolerhood, and children with special needs were referred to community day care centers. The research team members (primary care nurses) regularly visited these centers to monitor, consult on, and obtain clinical information about the children with special needs during preschoolerhood.
5. *School entry health check-up* Children at age 5 received a health check-up before school entry. For children with developmental concerns, detailed

interviews were conducted with the children and parents using an interview-based instrument, the Pervasive Developmental Disorders Autism Society Japan Rating Scale (Ito et al. 2012), and an IQ assessment was conducted by our research team.

Because diagnostic judgments by experienced clinicians are considered to be the “gold standard” for autism diagnosis (Volkmar et al. 2005), final diagnosis was decided according to DSM-IV-TR (American Psychiatric Association 2000) on the basis of all available information obtained after age 3 by the research team. IQs/DQs were assessed by different measures depending on mental age, using the Tanaka-Binet Intelligence Scale V for children, the Enjoji’s Analytical Developmental Test under age 4, or the Japanese version of the Wechsler Intelligence Scale for Children-Third Edition (WISC-III) at age 5.

Clinical measures were compared by group with the use of ANOVA and the Bonferroni multiple comparison test. The proportion of boys versus girls, developmental delay versus high-functioning, and the presence/absence of the targeted problems were compared with use of the Chi square test. Statistical analysis was performed using SPSS software. The protocol of this study was approved by the Ethics Committee of the National Center of Neurology and Psychiatry. This study was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments.

Results

Throughout the screening and surveillance process of the 1,851 children, we identified 51 children with ASD: 20 screen positives, 22 screen negatives, and 9 non responders (i.e., children who needed FUI but were missed among the attrition group) (Figs. 1, 2). Thirty-four children were directly evaluated by the research team (minimum ASD). Sixteen were diagnosed with autistic disorder (AD). Table 1 outlines their demographic and diagnostic characteristics. In addition, 17 children were clinically judged by the research team to have ASD on the basis of available information, such as that from local clinicians, which brought the total number of children with ASD up to 51 (maximum ASD).

Prevalence rate was estimated to be 0.0184 (95 % confidence interval [CI] 0.0123–0.0245), and 0.0276 (95 % CI 0.0201–0.03501) for minimum and maximum ASD, respectively. The boy/girl ratio of 2.8 and 2.2 and proportion of developmental delay of 38.2 and 52.9 % in the 34 and 51 children with ASD, respectively, were in parallel with the latest reported figures (Kim et al. 2011), indicating the representativeness of this sample. Regarding AD, the prevalence rate was estimated as 0.0086 (95 % CI 0.0044–0.0129).

The sensitivity, specificity, PPV, and likelihood ratio (LR) for maximum ASD, minimum ASD, and AD through both the first-stage screening and the entire two-stage screening are shown in Table 2. Calculations for the two-stage screening including FUI were based on 1,727 children after excluding 124 FUI non-responders. Re-screening with FUI improved the specificity, PPV, and LR but reduced the sensitivity for maximum and minimum ASD and AD. Since probability is influenced by prevalence of the disorder studied, we calculated the posttest probability assuming that a prevalence rate of 2.5 % for all ASDs according to Bayes’ theorem, giving a posttest probability of 0.47 and 0.51 for maximum and minimum ASD, respectively. These figures mean that almost one in every two screen positives will subsequently be diagnosed with ASD.

Among 319 screen positives at the first stage who needed FUI, only 195 were followed (response rate 61 %). One-hundred twenty-four non-responders (NR) had a significantly lower mean total M-CHAT-JV score (mean 2.81 ± 1.85) than the 195 responders (mean 3.35 ± 2.15) ($t = 2.32, p < 0.05$) and included significantly more girls (50 vs. 37 %) ($\chi^2 = 2.32, p < 0.05$), while neither group differed significantly in regard to age at M-CHAT-JV, critical items, or the proportion of nonverbal children at 18 months of age. Of the 124 NR, 9 were identified as having ASD before they were evaluated by our research team, 5 of whom had sought professional help regarding language delay.

The true positives (TP, $n = 20$), false positives (FP, $n = 24$), false negatives (FN, $n = 22$), and true negatives (TN, $n = 1661$) were compared according to demographic and diagnostic characteristics (Table 3). Although TP had significantly higher M-CHAT-JV total and critical scores than FP, FN, and TN ($ps < 0.001$), TP could not be discriminated from FP or FN by either sex ratio, maternal age at childbirth, perinatal problems, mother’s feeling of difficulty with child rearing at 18 months, or mother’s concerns about the child’s emotional or behavioral difficulties at 3 years. A comparison between TP and FN revealed that CARS, ADI-R, and ADOS scores at 3 years or older did not significantly differ between TP and FN, but there were significantly more children with developmental delay among TP (60 vs. 27 %, $p < 0.05$). As for the 24 FP cases, mothers of 22 children reported finding child-rearing difficult on the routine 18-month health check-up questionnaire, and those of 12 children expressed some concern about their child’s emotional or behavioral difficulties on the routine 3-year health check-up questionnaire. Although there were not necessarily objective records available to support their reports at or above 3 years of age, one boy had a DQ of 61 at 2, and 3 boys were clinically judged as having mild developmental delay at the 3-year pediatric check-up. In addition, the research team evaluations

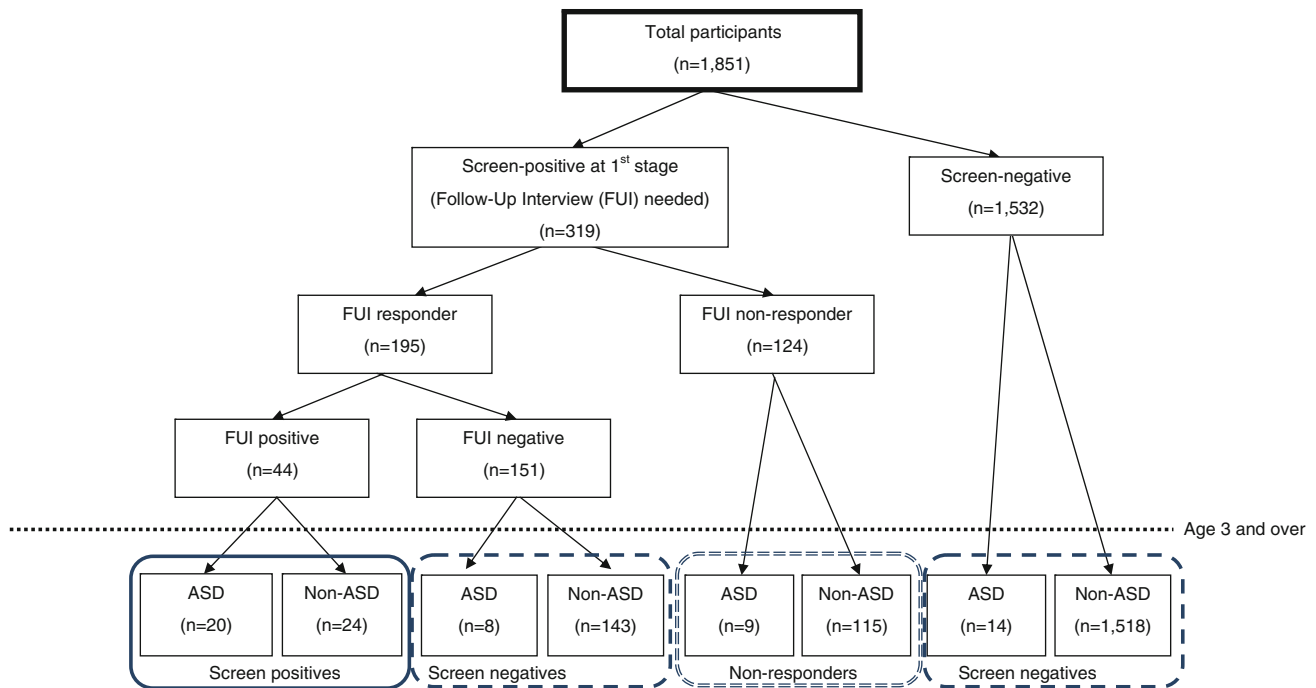


Fig. 2 Results of screening. Non-responders are children who needed a follow-up telephone interview but were missed among the attrition group

Table 2 Psychometric properties of the M-CHAT-JV screening

		Sensitivity 95 % CI	Specificity 95 % CI	PPV 95 % CI	Likelihood ratio 95 % CI
Maximum ASD	1st	0.725 (37/51)	0.843 (1,518/1,800)	0.116 (37/319)	4.631
		0.594–0.828	0.840–0.846	0.095–0.132	3.703–5.382
	1st + 2nd	0.476 (20/42)	0.986 (1,661/1,685)	0.455 (20/44)	33.433
		0.351–0.596	0.983–0.989	0.335–0.569	20.228–52.908
Minimum ASD	1st	0.765 (26/34)	0.839 (1,524/1,817)	0.082 (26/319)	4.742
		0.603–0.875	0.836–0.841	0.064–0.093	3.670–5.497
	1st + 2nd	0.613 (19/31)	0.985 (1,671/1,696)	0.432 (19/44)	41.579
		0.457–0.748	0.982–0.988	0.322–0.527	25.967–60.921
AD	1st	0.625 (10/16)	0.832 (1,526/1,835)	0.031 (10/309)	3.712
		0.388–0.815	0.830–0.833	0.019–0.041	2.276–4.885
	1st + 2nd	0.500 (7/14)	0.978 (1,676/1,713)	0.159 (7/44)	23.149
		0.274–0.725	0.977–0.980	0.087–0.231	11.695–36.670

“Maximum ASD” referred to 51 children who were classified as autism spectrum disorder (ASD) based on available information

“Minimum ASD” referred to 34 children who were directly evaluated and diagnosed as ASD by the research team. “AD” referred to 16 children who were directly evaluated and diagnosed as autistic disorder by the research team

PPV positive predictive value

confirmed two subthreshold ASD cases: one girl was diagnosed with ASD at both age 2 (IQ 68) and 3 (IQ 89), but at age 4 (IQ 123) the symptoms no longer met the diagnostic criteria. Another boy was a floppy infant with autistic features at age 2, and subsequently motor developmental delay became apparent with reduced autistic symptoms.

Discussion

This study aimed to examine prospectively the utility of an autism-specific screening in conjunction with community developmental surveillance for a non-selected Japanese population. Two-stage screening with the M-CHAT-JV identified 20 of 51 children with ASD across all intellectual

Table 3 Comparison of demographic and diagnostic characteristics: true positive, false positive, false negative, and true negative

	True positive		False positive		False negative		True negative		F	χ^2	p
	n	M (SD)	n	M (SD)	n	M (SD)	n	M (SD)			
Sex ratio (M : F)	20	14 : 6	24	16 : 8	22	15 : 7	1,661	835 : 826		8.2	<0.05
Age at M-CHAT-JV (months)	20	18.6 (0.5)	24	18.5 (0.5)	22	18.7 (0.7)	1,661	18.7 (0.6)	0.5	0.1	>0.05 ^a
M-CHAT-JV total	20	6.3 (3.0)	24	4.5 (2.0)	22	1.7 (1.6)	1,661	0.8 (1.0)	238.0		<0.001
M-CHAT-JV critical10	20	3.5 (2.1)	24	2.5 (1.2)	22	0.8 (1.2)	1,661	0.1 (0.5)	322.0		<0.001
Gestational age (weeks)	18	38.2 (2.1)	23	39.0 (1.7)	22	39.1 (1.4)	1,604	39.0 (1.6)	1.7		>0.05
Birth weight (g)	18	2949.5 (517.7)	23	3130.3 (389.5)	22	3029.6 (425.2)	1,647	3049.3 (419.4)	0.9		>0.05
Maternal age at childbirth	20	30.2 (5.1)	24	30.8 (4.6)	22	32.8 (4.1)	1,653	30.1 (4.7)	1.8		>0.05
Perinatal problems (present: absent)	19	4 : 15	24	6 : 18	22	6 : 14	1,597	426 : 1171		0.5	>0.05
Mother's feeling of difficulty with child-rearing at 18 months (present: absent)	20	19 : 1	23	22 : 1	22	21 : 1	1,659	1633 : 26		3.5	>0.05
Mother's concerns about the child's emotional or behavioral difficulties at 3 years (present: absent)	14	8 : 6	20	12 : 8	22	7 : 15	1,646	1120 : 526		14.2	<0.001
CARS total	20	34.2 (5.7)	10	23.7 (5.1)	12	32.8 (4.0)	4	26.4 (1.9)	14.5		<0.001 ^b
ADI-R toddler total	17	26.9 (9.0)	6	16.5 (7.1)	12	23.0 (4.9)	4	15.5 (9.8)	4.3		<0.05 ^b
ADOS (a) + (b) total	9	13.7 (3.7)	2	1.0 (0.0)	10	12.7 (3.9)	3	6.7 (3.8)	9.9		<0.01 ^b
IQ/DQ	20	73.3 (27.7)	11	91.1 (15.6)	18	88.3 (26.0)	5	78.2 (6.7)	0.9		>0.05
≥85	6		8		11		1				
70–84	2		2		3		4				
50–69	9		1		1		0				
35–49	2		0		3		0				
<35	1		0		0		0				
Developmental delay (%)	12/20 (60 %)		4/24 (17 %)		6/22 (27 %)					8.9	<0.005
										4.6	<0.05

^a Three groups of TP, FN, and FP were compared using the Chi square test

^b Three groups of TP, FN, and FP were compared using ANOVA

functioning levels. This indicates that the autism-specific screening at 18 months of age in primary health settings is feasible and useful when combined with community-based surveillance for preschoolers.

The controversial issue regarding the age of screening was partly answered in this study. Our findings indicate that the age of 18 months can be applied with acceptable predictive values, better than those in the earlier pioneering work (Baird et al. 2000). A possible explanation for why the M-CHAT-JV screening could identify children with ASD at this age is that the M-CHAT items might represent age-specific social development such as joint attention and pretend play that few typically developing children lack at 18 months (Inada et al. 2010; Oosterling et al. 2010), and that it could detect nonverbal social maldevelopment even in children with high-functioning ASD (HFASD). In the present study, only 30 % of 20 detected children with ASD had IQ at or above 85 and the 60 % had IQ/DQ below 70 (see Table 3). We found that the proportion of children with IQ/DQ below 70 was significantly greater among true-positive children than false-negative children, although the severity of autistic symptoms assessed by the CARS, ADI-R, or ADOS at 3 years did not differ between them. This finding suggests that the parent-report M-CHAT-JV screening measure at 18 months was more sensitive to low-functioning ASD than to high-functioning ASD, similar to earlier studies with unselected/low-risk children (Pandey et al. 2008; Kleinman et al. 2008; Baron-Cohen et al. 1996) in which detected children were mainly developmentally retarded. If the reduced sensitivity to high-functioning ASD is partly due to a lack of parental awareness, in addition to the parent-report M-CHAT-JV questionnaire, it could be possible to improve sensitivity by direct observation of some of its items by primary health nurses. In order to examine this hypothesis, a prospective study is currently underway to compare the sensitivity of the parent-report M-CHAT alone with that of the M-CHAT plus direct observation.

We recognize that we could not evaluate all screen-positive children directly, but we did instead clinically judge children who were not directly evaluated based on the information available from community surveillance. Since early detection of ASD should be economically balanced with existing surveillance procedures (Charman et al. 2002), in the absence of any better alternative screen, we recommend enhancing community developmental surveillance by supplementing it with the M-CHAT screen-rescreen procedure. Although a one-point screening model may be cost-effective, we conclude that a comprehensive model comprising repetitive screening and subsequent community surveillance will be more appropriate, considering the various developmental trajectories of children with ASD (Fernell et al. 2010; Robins et al. 2001). An

advantage of the time lag associated with the screen-rescreen procedure might be that it gives parents time to pay attention to their child's ongoing social development. To answer definitively the issue about the optimal age of screening, more empirical studies are needed and the merits and demerits for each screening procedure should be determined based on long-term follow-up data.

Our results indicated that there were at least twice as many children with HFASD missed ($n = 8$) as those detected ($n = 16$) at screening, which is consistent with Kleinman et al. (2008). In general, parents seem to be unaware of reduced social development in their child with HFASD. However, there is the possibility that these missed children show a different developmental trajectory in the very early years from that of the detected children.

Many clinicians will likely be concerned at the high screen-positive rate at the first stage of screening (17 %) because parents of children who were incorrectly suspected of having ASD might suffer unnecessary distress. This high rate might be related to the high attrition rate of 39 % (124/319) between the two stages. Since we could not systematically investigate the attrition group (the non-responders), details of the referral pattern for children with ASD who were screen positive at the first stage but who were later missed are not clear in this study. If we raise the first-stage screening threshold to approach the original one (any 3 from the total 23 or any 2 from the critical set criteria), this reduces screen positive cases ($n = 39$), and as a result slightly increases the PPV from 0.455 up to 0.462 (18/39) but also reduces the sensitivity from 0.476 up to 0.439. Closer inspection reveals that mothers of the majority of the false-positive children actually had been concerned about their child-rearing by age 3, and through evaluations, several children were confirmed to have problems in either cognitive, language, social, or motor domains even though the symptoms did not meet the diagnostic criteria for ASD. These findings could suggest that the false-positive cases in our study might have neurodevelopmental symptoms that extend beyond those of ASD, which are in common with those seen in many children referred to clinics (Gillberg 2010). Following this thought further, the M-CHAT screening at 18 months may be sensitive to children with mild but overlapping neurodevelopmental problems in multiple domains to some degree. This issue should be investigated in future studies using a comprehensive neurodevelopmental assessment tool.

Two major limitations exist in the current study. First, although efforts were made in cooperation with local day nurseries and clinicians to identify missed screen-positive and ASD-suspected screen-negative cases, the attrition rate was high and community-based developmental surveillance was not then sufficient in itself to monitor all children. The final diagnosis of 17 ASD cases was made based on such

indirect information. There is also the possibility that we missed a subset of children with ASD, particularly those with milder autistic symptoms, average intelligence, or girls, for whom diagnosis of ASD tends to be delayed (Mandell et al. 2005; Shattuck et al. 2009). As a result, the sensitivity and specificity of the M-CHAT-JV that we calculated based on these results can only be considered estimates of their upper bounds. Second, although various standardized instruments were used for case ascertainment of strictly defined ASD cases, the most standard ones such as the ADOS and ADI-R were not available in Japan at the beginning of this study. The total prevalence rate in our study is similar to the latest figure available from a study using strict scientific methodology (Charman et al. 2002), which indicates the quality of case ascertainment in our study.

In summary, two-stage autism-specific screening using the M-CHAT with some modification of the threshold could effectively identify Japanese children with ASD, even HFASD. We would like to emphasize that not only screening but also continual community-based developmental surveillance is necessary for detecting children with ASD. Such enhancement of multidisciplinary community assessment should result in promoting the development of children with ASD and improve their quality of life (Kamio et al. 2013).

Acknowledgments This work was funded by RISTEX (Japan Science and Technology Agency) of Japan. We would like to thank Drs. Deborah Fein and Diana Robins for their invaluable advice on this project and comments about the findings, and Dr. Hisateru Tachimori for statistical advice. The contents of this paper were presented at the “Exploring Autism Research Collaboration between Japan and United States Joint Academic Conference on Autism Spectrum Disorders” held in Tokyo, Japan, December 1–3, 2011.

Conflict of interest The authors declare that they have no conflict of interest with respect to this article.

References

- Allen, C. W., Silove, N., Williams, K., & Hutchins, P. (2007). Validity of the social communication questionnaire in assessing risk of autism in preschool children with developmental problems. *Journal of Autism and Developmental Disorders*, *37*, 1272–1278.
- American Psychiatric Association (2000) *Diagnostic and statistical manual of mental disorders*. 4th ed., text revision. (DSM-IV-TR). Washington, DC: American Psychiatric Association.
- Baird, G., Charman, T., Baron-Cohen, S., Cox, A., Swettenham, J., Wheelwright, S., et al. (2000). A screening instrument for autism at 18 months of age: A 6-year follow-up study. *Journal of American Academy of Child and Adolescent Psychiatry*, *39*, 694–702.
- Baron-Cohen, S., Allen, J., & Gillberg, C. (1992). Can autism be detected at 18 months? The needle, the haystack, and the CHAT. *The British Journal of Psychiatry*, *161*, 839–843.
- Baron-Cohen, S., Cox, A., Baird, G., Swettenham, J., Nightingale, N., Morgan, K., et al. (1996). Psychological markers in the detection of autism in infancy in a large population. *The British Journal of Psychiatry*, *168*, 158–163.
- Barton, M., Chlebowski, C., & Fein, D. (2008). Screening for autism spectrum disorders in young children: A review. *Seishin Hoken Kenkyu (Journal of Mental Health)*, *54*, 7–28.
- Charman, T., Baron-Cohen, S., Baird, G., Cox, A., Swettenham, J., Wheelwright, S., et al. (2002). Is 18 months too early for the CHAT? *Journal of American Academy of Child and Adolescent Psychiatry*, *41*, 235–236.
- Charman, T., Baron-Cohen, S., Baird, G., Cox, A., Wheelwright, S., Swettenham, J., et al. (2001). Commentary: The modified checklist for autism in toddlers. *Journal of Autism and Developmental Disorders*, *31*, 145–148.
- Chlebowski, C., Robins, D. L., Barton, M. L., & Fein, D. (2013). Large-scale use of the modified checklist for autism in low-risk toddlers. *Pediatrics*, *131*, e1121–e1127. doi:10.1542/peds.2012-1525.
- Dietz, C., Swinkels, S., van Daalen, E., van Engeland, H., & Buitelaar, J. K. (2006). Screening for autistic spectrum disorder in children aged 14–15 months. II: Population screening with the early screening of autistic traits questionnaire (ESAT). Design and general findings. *Journal of Autism and Developmental Disorders*, *36*, 713–722.
- Eaves, L. C., Wingert, H., & Ho, H. H. (2006). Screening for autism: Agreement with diagnosis. *Autism*, *10*, 229–242.
- Fernell, E., Hedvall, A., Norrelgen, F., Eriksson, M., Höglund-Carlsson, L., Barnevik-Olsson, M., et al. (2010). Developmental profiles in preschool children with autism spectrum disorders referred for intervention. *Research in Developmental Disabilities*, *31*, 790–799.
- Gillberg, C. (2010). The ESSENCE in child psychiatry: Early symptomatic syndromes eliciting neurodevelopmental clinical examinations. *Research in Developmental Disabilities*, *31*, 1543–1551.
- Gray, K. M., Tonge, B. J., Sweeney, D. J., & Einfeld, S. L. (2008). Screening for autism in young children with developmental delay: An evaluation of the developmental behaviour checklist: Early screen. *Journal of Autism and Developmental Disorders*, *38*, 1003–1010.
- Inada, N., Kamio, Y., & Koyama, T. (2010). Developmental chronology of preverbal social behaviors in infancy using the M-CHAT: Baseline for early detection of atypical social development. *Research in Autism Spectrum Disorders*, *4*, 605–611.
- Inada, N., Koyama, T., Inokuchi, E., Kuroda, M., & Kamio, Y. (2011). Reliability and validity of the Japanese version of the modified checklist for autism in toddlers (M-CHAT). *Research in Autism Spectrum Disorders*, *5*, 330–336.
- Ito, H., Tani, I., Yukihiro, R., Adachi, J., Hara, K., Ogasawara, M., et al. (2012). Validation of an interview-based rating scale developed in Japan for pervasive developmental disorders. *Research in Autism Spectrum Disorders*, *6*, 1265–1272.
- Johnson, C. P., & Myers, S. M. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics*, *120*, 1183–1215.
- Kamio, Y., Inada, N., & Koyama, T. (2013). A nationwide survey on quality of life and associated factors of adults with high-functioning autism spectrum disorders. *Autism*, *17*, 16–27.
- Kim, Y. S., Leventhal, B. L., Koh, Y. J., Fombonne, E., Laska, E., Lim, E. C., et al. (2011). Prevalence of autism spectrum disorders in a total population sample. *The American Journal of Psychiatry*, *168*, 904–912.
- Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H. C., Esser, E. L., et al. (2008). The modified checklist for

- autism in toddlers: A follow-up study investigating the early detection of autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 38, 827–839.
- Kurita, H., Miyake, Y., & Katsuno, K. (1989). Reliability and validity of the childhood autism rating scale-Tokyo version (CARS-TV). *Journal of Autism and Developmental Disorders*, 19, 389–396.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H, Jr, Leventhal, B. L., DiLavore, P. C., et al. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30, 205–223.
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24, 659–685.
- Mandell, D. S., Novak, M. M., & Zubritsky, C. D. (2005). Factors associated with age of diagnosis among children with autism spectrum. *Pediatrics*, 116, 1480–1486.
- Mothers' & Children's Health & Welfare Association. (2007). *Maternal and child health statistics of Japan (in Japanese)*. Tokyo.
- Oosterling, I. J., Wensing, M., Swinkels, S. H., van der Gaag, R. J., Visser, J. C., Woudenberg, T., et al. (2010). Advancing early detection of autism spectrum disorder by applying an integrated two-stage screening approach. *Journal of Child Psychology and Psychiatry*, 51, 250–258.
- Pandey, J., Verbalis, A., Robins, D. L., Boorstein, H., Klin, A. M., Babitz, T., et al. (2008). Screening for autism in older and younger toddlers with the modified checklist for autism in toddlers. *Autism*, 12, 513–535.
- Robins, D. L. (2008). Screening for autism spectrum disorders in primary care settings. *Autism*, 12, 537–556.
- Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The modified checklist for autism in toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 31, 131–144.
- Rogers, S. J., & Vismara, L. A. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child and Adolescent Psychology*, 37, 8–38.
- Schopler, E., Reichler, R. J., & Renner, B. R. (1988). *The childhood autism rating scale (CARS)*. Los Angeles: Western Psychological Services.
- Shattuck, P. T., Durkin, M. D., Maenner, M., Newschaffer, C., Mandell, D. S., Wiggins, L., et al. (2009). The timing of identification among children with an autism spectrum disorder: Findings from a population-based surveillance study. *Journal of American Academy of Child and Adolescent Psychiatry*, 48, 474–483.
- Tsuchiya, K. J., Matsumoto, K., Yagi, A., Inada, N., Kuroda, M., Inokuchi, E., et al. (2012). Reliability and validity of autism diagnostic interview-revised, Japanese version. *Journal of Autism and Developmental Disorders*,. doi:10.1007/s10803-012-1606-9.
- Volkmar, F., Chawarska, K., & Klin, A. (2005). Autism in infancy and early childhood. *Annual Review of Psychology*, 56, 315–336.